Mapping and characterization of the *required to maintain repression10* locus affecting paramutation

Research Thesis

Presented in partial fulfillment of the requirements for graduation with research distinction in Molecular Genetics the undergraduate colleges of The Ohio State University

by

Emily McCormic

The Ohio State University

April 2019

Project Advisor: Dr. Jay Hollick, Department of Molecular Genetics

Table of Contents

Abstract

Paramutation describes a poorly understood biological process in which one allele of a gene causes heritable changes to the corresponding allele. Such changes seen at multiple loci in diverse organisms lead to inheritance patterns that appear to defy the first law of Mendelian genetics. The underlying molecular mechanism responsible for paramutation is best understood in corn (*Zea mays*) where paramutation occurring at the *Pl1-Rhoades (Pl1-Rh)* allele of the *purple plant1 (pl1)* locus causes meiotically heritable changes in gene regulation. Plants with an unrepressed *Pl1-Rh* allele exhibit dark purple coloration, while plants with a repressed paramutant derivative (denoted *Pl'*) lack such coloration. The suppression of *Pl1-Rh* can be attributed to functions encoded by *required to maintain repression (rmr)* genes. Five different *rmr* genes identified by mutations have been previously reported. Identifying all the individual *rmr* gene products is critical to understanding the molecular mechanism responsible for the paramutation process.

Having determined that two mutations define a novel *rmr* locus designated *rmr10*, I aimed to characterize the function and identify the molecular nature of this gene. I first compared *Pl1-Rh* mRNA levels between mutant and non-mutant siblings and between reference *Pl-Rh* and *Pl'* plants to test the hypothesis that the *rmr10* locus affects RNA abundance.

To identify the molecule encoded by the *rmr10* locus, with the help of collaborators, I mapped the position of the *rmr10* locus in the genome and then used a candidate gene approach to evaluate possible gene models using an open-source bioinformatics pipeline. Likely gene candidates were selected based on relevance to the biological processes of DNA/histone modifications or small RNA production.

Introduction

The study of gene regulation is crucial in understanding how genetics can create diverse and complex organisms which define life on earth. Gene expression goes beyond the sequence of DNA bases contained in a cell's nucleus- myriad systems of regulation exist to affect the ways in which the genetic code produces phenotypes. Epigenetics describes heritable changes conferred by mechanisms which do not affect the DNA sequence.[1] One such epigenetic phenomenon that has expanded geneticists' understanding of gene regulation is paramutation. Paramutation was first described in plants to explain an unusual inheritance behavior that seemingly defied Mendel's First Law of Inheritance.[2,3] In a heterozygous state, one allele of a gene can alter the expression of the corresponding allele in an epigenetic and often reversible fashion.[4] Paramutation has been previously observed at multiple loci in *Zea mays,* including *r1* and *b1*.[5] Our lab has studied this phenomenon at the *purple plant1* (*pl1*) locus in *Zea mays* where the *Pl1-Rhoades (Pl1-Rh)* allele, normally causing dark coloration in the plant's anthers, can be epigenetically repressed in the *Pl'* state and exhibit light coloration in the anthers (Figure 1). The *pl1* locus encodes a transcription factor that is involved in the production of the anthocyanin pigment responsible for coloration in different parts of the plant, including the anthers.[6]

Regulation of the *Pl1-Rh* allele can cause multiple levels of repression rated on a scale of anther color score (ACS) 1-7 with 1 lacking purple coloration, and 7 presenting full coloration (Figure 2).[7]

Repression of the *Pl1-Rh* allele is maintained by the protein products encoded from *required to maintain repression* (*rmr*) genes, five of which have been previously reported. These genes were identified in an ethyl methanesulfonate (ems) screen and were determined to define novel loci based on genetic complementation testing. All the proteins encoded by these genes

function in a small RNA (sRNA) biogenesis pathway that may direct epigenetic changes to the DNA (Figure 3).[5]

I determined one such mutation created via ems mutagenesis, ems073240, defined a locus unique to previously discovered *rmr* genes based on complementation testing and molecular characterization. The ems073240 mutation was then designated the reference allele of a novel *rmr* locus, *rmr10*.

Plants homozygous for the mutant *rmr10* allele exhibit darker anther coloration, suggesting RMR10 is involved in paramutation at the *pl1* locus by maintaining the repression of the *Pl'* state of the *Pl1-Rh* allele. Because the precise function of RMR10 is unknown, I first characterized the *Pl1-Rh* RNA abundance in *rmr10* mutants and non-mutant siblings with the expectation that RNA abundance would be higher in mutants given their darker coloration. Previous experiments, including an RNase Protection Assay, also supported the hypothesis that RNA abundance would be higher in a *rmr* mutant exhibiting the *Pl-Rh* phenotype.[6]

With the help of collaborators, I then mapped the location of the *rmr10* locus in the genome using RNA-seq data from a pool of *rmr10* mutants using a bioinformatics pipeline that could narrow down candidate regions in the genome and predict how the single nucleotide polymorphisms (SNPs) generated from the EMS mutagenesis would affect the protein products of the gene models within the candidate regions.

Figure 1: The two epigenetic states of the *Pl1-Rh* allele. Both are genetically identical states, but the *Pl'* state (right) is epigenetically repressed and features reduced anther coloration as compared to the reference *Pl-Rh* state (left).

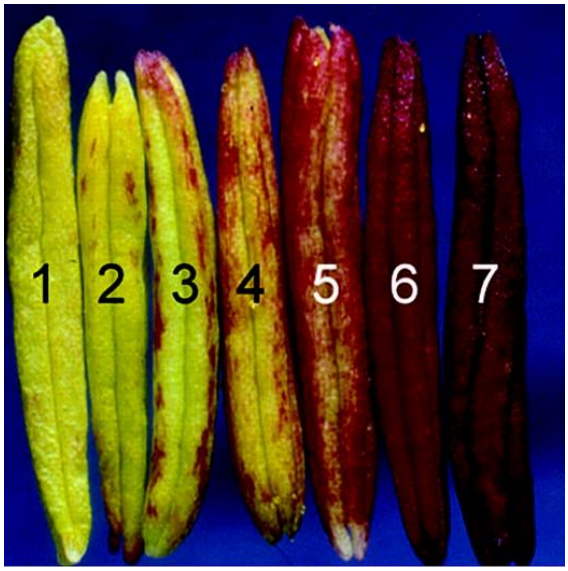

Figure 2: The anther color score (ACS) system used to quantify anther color phenotype. ACS 1-4 represents the *Pl'* phenotypic state while ACS 5-7 represents the *Pl-Rh* state.

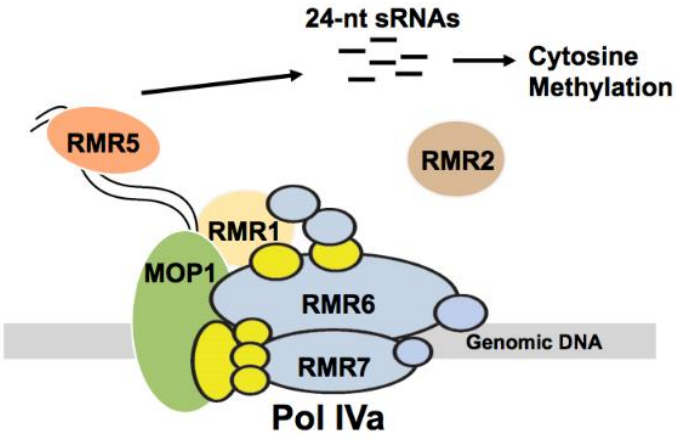

Figure 3: The roles of some *rmr* factors in a small RNA biogenesis pathway affecting cytosine methylation. *rmr6* encodes the largest subunit of RNA Polymerase IV. *rmr7* is one of the second largest subunits of RNA Polymerase IV. *rmr1* is an SNF2-like RNA helicase. *rmr5* encodes an endonuclease that cleaves the double-stranded RNA produced by *mop1*, an RNA-dependent RNA polymerase. The specific function of *rmr2* function in the pathway is currently unknown.[8-14]

Materials and Methods

Synthesis of the *rmr10* Mutant Material

The generation of the B73, A619, and A632 stocks introgressed with *Pl1-Rh* and EMS mutagenesis protocols has been previously described.[15] The ems073240 mutation was backcrossed into the A632 line four times (Figure 4). This backcross strategy allowed for the comparison of RNA-seq data from plants with the ems073240 mutation against reference genomes to determine candidate regions.

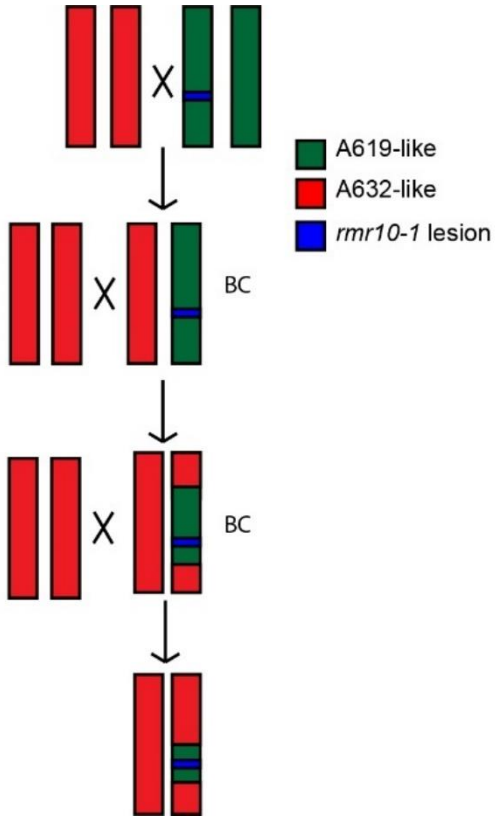

Figure 4: ems073240 backcross strategy. Chromosomes are color-coded based on inbred line character, as determined by single nucleotide polymorphisms (SNPs) unique to each line. After each backcross, plants with the mutant allele are selected for and backcrossed again. The figure represents two of the four total backcrosses. Since the ems073240 lesion was induced in the A619 background and introgressed into A632 to BC4F2, the mutant lesion should be in a region where A619 SNPs are concentrated. A619-like regions should account for less than 10% of the genome.

Confirmation of the *rmr10* Allele

The ems073240 mutation was determined to define a unique mutant locus based on complementation testing with other known *rmr* mutations (Table 1), apart from *mediator of paramutation1* (*mop1*). Plants homozygous for the ems073240 mutation were tested with a series of derived cleaved amplified polymorphic sequence (dCAPS) markers around the *mop1* locus to determine the A619 or A632 character of the sequence, based on the expectation that plants with ems073240 would have A619-like sequence surrounding the *mop1* locus if the ems073240 and *mop1* lesions defined the same locus (Figure 5). The dCAPS markers (Table S1) flanked the region around *mop1*, including a marker 6 Mb away from the locus. Genomic DNA from mutant florets was isolated and amplified in a PCR reaction (Table S2). The markers were designed to amplify a single nucleotide difference in the A619-like or A632-like sequence which could then

be differentially cut with the restriction enzyme *Dpn*I. The restriction digest product was then run out on a 2% agarose gel (150V for 20 minutes) and stained with ethidium bromide (10 minutes).

Parental Genotypes		Progeny		
Male (ems allele)	Female	Identifier	Anther Phenotype	
			Pl-Rh	*Pl'*
073240	+/*rmr1-1*	111387	0	15
073240	+/*rmr2-1*	112665	0	15
073240	+/*rmr3-1*	142230	0	22
073240	+/*rmr5-1*	112286	0	12
073240	+/*rmr6-1*	112259	0	12
073240	+/*rmr7-1*	112524	0	17
073240	+/*rmr8-1*	112283	0	15
073240	+/*rmr12-1*	112269	0	13
073240	+/062986	112272	5	6
073240	+/071801	112293	0	16
073240	+/072093	112686	0	11
073240	+/073300	141385	0	20

Table 1: Complementation testing data for ems073240. The mutation complemented all known *rmr* factors apart from *mop1* (for which no complementation testing was done). ems062986 shows possible complementation and was designated *rmr10-2*.

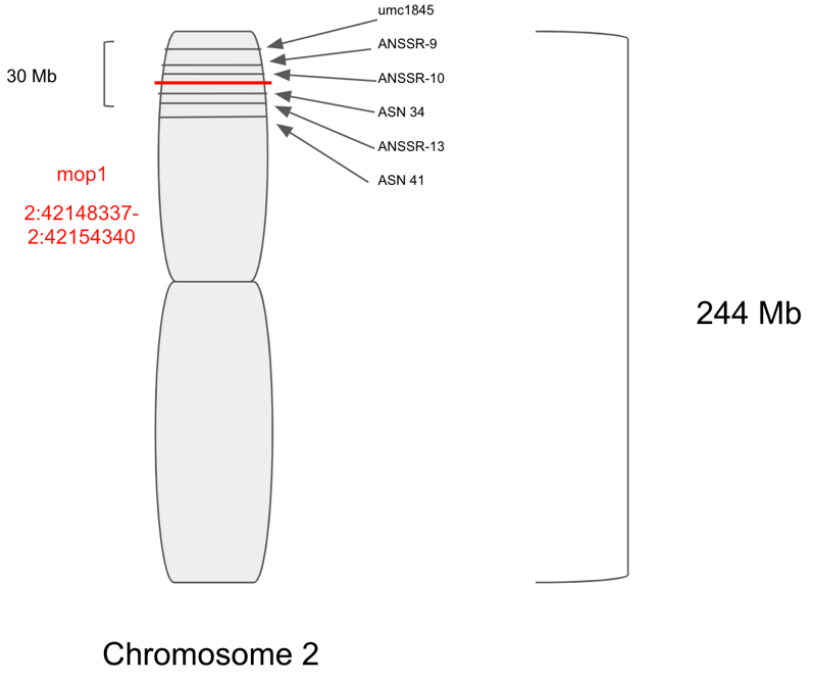

Figure 5: Molecular characterization of the region surrounding the *mop1* locus. Six markers in a 30 Mb region were used to amplify sequences in the region which could be digested with a restriction enzyme and differentially cut based on A619 or A632 character.

Characterization of RNA Abundance

The relative abundance of RNA produced from the *Pl1-Rh* allele was measured using a quantitative real time PCR (qRT-PCR) assay. RNA was isolated from the anthers of three mutant and three non-mutant plants using a TRIzol extraction and converted to cDNA. The qRT-PCR assay was set up by combining each cDNA sample with a reaction mixture containing SYBR Sensi-mix and primers designed to amplify the *Pl1-Rh* allele (protocols previously described).[15] Each of the three biological replicates were tested three times, creating technical and biological triplicates. The samples were also amplified with primers designed to amplify the *alanine aminotransferase5* (*aat*) gene as a control, with the expectation that all plants have a consistent expression of *aat* to which *Pl1-Rh* mRNA levels can be normalized.

I first tested *rmr10* mutant and non-mutant siblings with a primer set that was previously successful in amplifying the *pl1* locus, normalized to levels of *aat*. I then tested wild-type, true-breeding *Pl-Rh* and *Pl'* A619 anthers with the qRT-PCR assay to determine if the assay was able to detect differences in the abundance of *Pl1-Rh* RNA with a separate primer set designed to amplify *Pl1-Rh* mRNA (Figure 6 and Table S3).

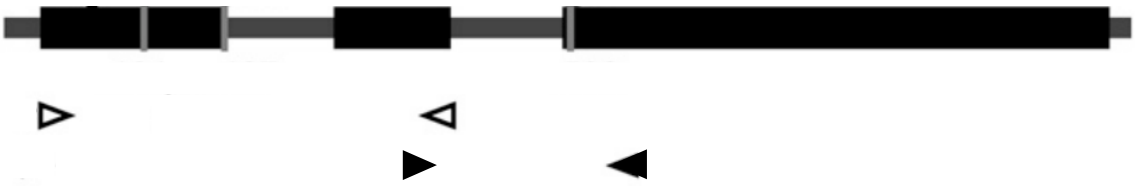

Figure 6: *Pl1-Rh* gene model with primers used for qRT-PCR analyses. Open arrows represent the primers used to amplify *Pl1-Rh* in A619 controls, solid arrows represent the primers used to amplify *Pl1-Rh* in plants homozygous for the mutant *rmr10* allele.[16]

Mapping of the *rmr10* Locus

RNA was isolated from the florets of plants homozygous for the mutant *rmr10-1* allele. These RNA isolations were adjusted to the same concentration of 60 ng/μL and combined to form a pool containing 16 samples. I sent the RNA pool to Novogene to create a Eukaryotic Strand-Specific Transcriptome Library and generate an RNA-seq data set. Jay Hollick then trimmed the raw data (127,223,896 reads) for quality control to produce 126,866,324 total reads (116,671,928 non-rRNA and non-tRNA reads). Our collaborator, Michael Sovic, then acquired HapMap data for the A619 and A632 lines to gather variants in A619 and A632 as compared to the reference B73 genome. The HapMap data was examined for informative sites that were homozygous in either A619 or A632, but not in both. The RNA-seq data was then compared with these informative areas to identify A619- and A632-like variants shared with the RNA-seq data set. The sequence was divided into 1 Mb bins and the proportion of A619-like variants within each section was calculated, with a minimum requirement of at least ten sequence reads representing a given variant to ensure variation in the sequence was due to a polymorphism rather than a lack of sequence coverage. Variants unique to the RNA-seq data within these informative A619-like areas were then identified by comparison to known SNPs in all available haplotype data. SNPs not represented in any inbred line or haplotype were considered unique to the RNA-seq data set. This allowed for the manual for the manual inspection of the gene models within the candidate regions to ensure sequence coverage and expression in the RNA-seq data. Jay Hollick identified gene models and their presumed biological function in these areas, and I narrowed down candidates based on relevant function and sufficient read coverage in the RNA-seq data set. Jay Hollick then used the snpEff tool (http://snpeff.sourceforge.net) to determine which gene models contained SNPs that could disrupt protein function. [17-19]

Results

Confirmation of the *rmr10* Mutation

Amplified and digested genomic DNA from homozygous *rmr10-2* mutants had a banding pattern consistent with the A632 genomic DNA control, suggesting the area surrounding the *mop1* lesion was not A619-like and therefore, did not contain the mutant lesion (Figure 7). Given the complementation to other known *rmr* factors, a molecular character that suggested ems073240 did not define the *mop1* locus, and non-complementation to a possible second mutant allele defined by ems062986, the mutant allele defined by ems073240 was designated *rmr10-1*.

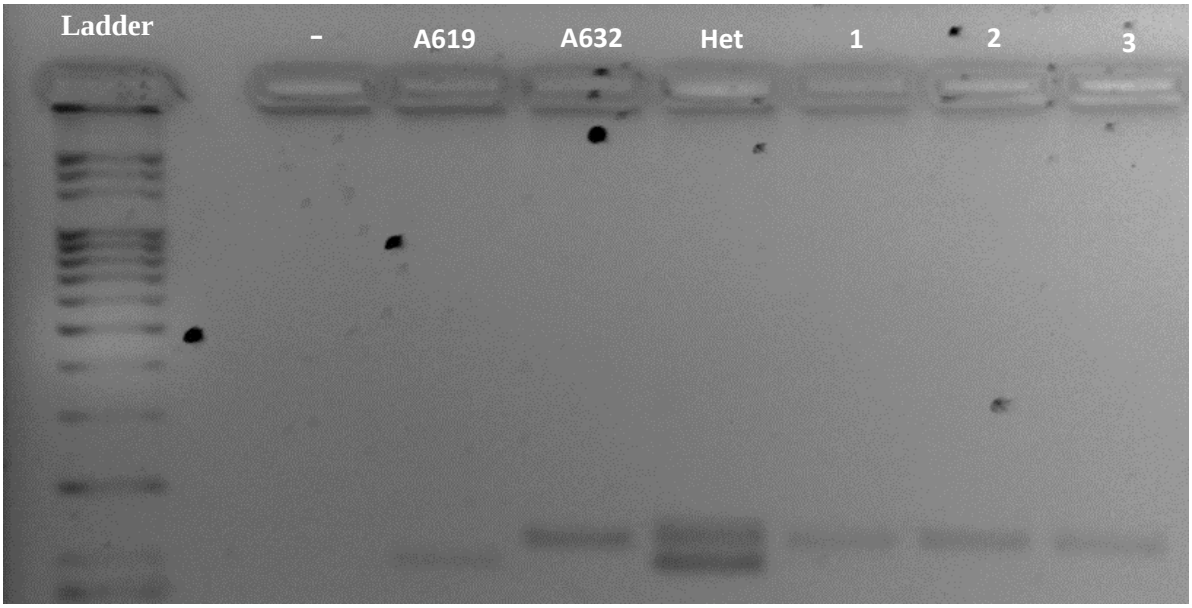

Figure 7: Gel electrophoresis results confirming the A632-like character around the *mop1* locus, consistent with the conclusion that ems073240 did not define *mop1* and represents a novel locus. The lanes from left to right: DNA ladder, negative control (water), A619 genomic DNA control, A632 genomic DNA control, A619/A632 heterozygote genomic DNA control, ems073240 DNA samples 1-3.

Characterization of RNA Abundance

RNA levels in A619 *Pl-Rh* and *Pl'* controls and *rmr10-1* mutant and non-mutant siblings were quantified with a quantitative real-time PCR assay. The A619 controls showed a significant difference in *Pl1-Rh* mRNA levels (H^o= No difference in *Pl1-Rh* mRNA abundance, H^a=Difference in *Pl-Rh* mRNA abundance, p=0.0031, two-sample *t*-test) (Figure 8). The *rmr10* mutant and non-mutant siblings did not show a significant difference in *Pl1-Rh* mRNA levels (Ho= No difference in *Pl1-Rh* mRNA abundance, H_a=Difference in *Pl-Rh* mRNA abundance *p*=0.8659, two-sample *t*-test) (Figure 9).

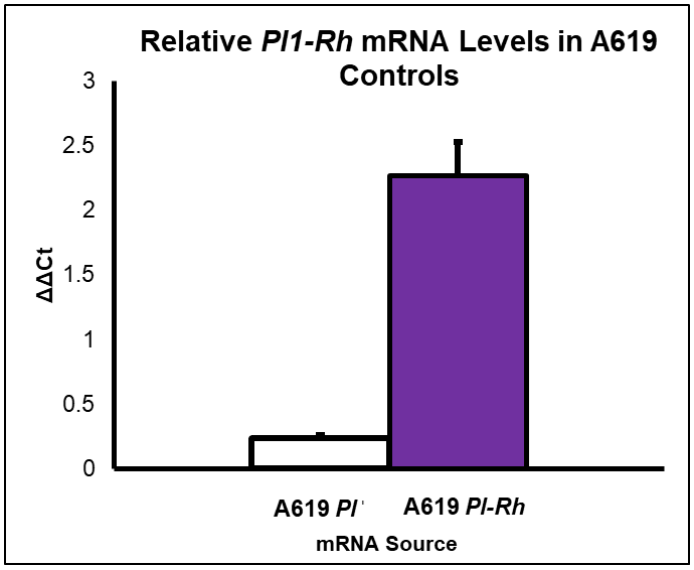

Figure 8: Relative *Pl1-Rh* mRNA level of wild-type A619 *Pl-Rh* and *Pl'* plants as measured by qRT-PCR. *p*=0.0031, two-sample *t*-test. Values based on biological and technical triplicates.

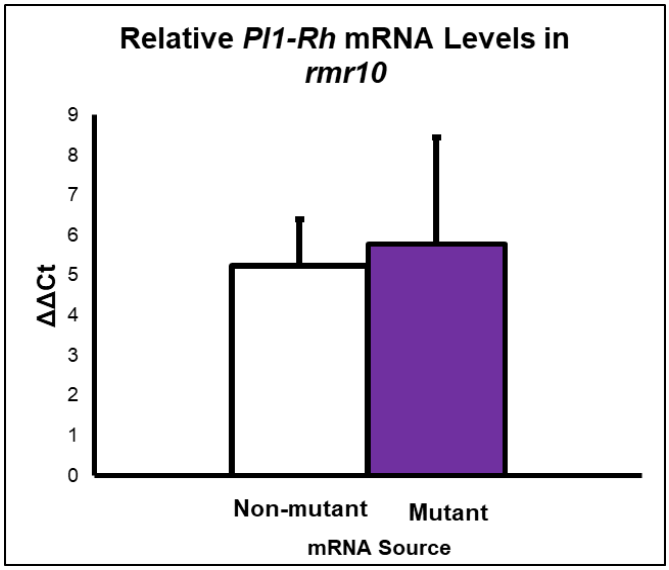

Figure 9: Relative *Pl1-Rh* mRNA level of mutant and non-mutant *rmr10* siblings as measured by qRT-PCR. *p*=0.8659, two-sample *t*-test. Values based on biological and technical triplicates.

Mapping of the *rmr10* Locus

The analysis of the *rmr10* RNA-seq data set to reference A619, A632, and B73 genomes showed a strong bias for A619-like SNPs in regions on chromosomes *1*, *2*, *4*, and *6*. Candidate regions have a higher proportion of A619-like SNPs compared to A632-like SNPs and have enough read coverage to ensure there is enough information to determine if the SNPs are A619- or A632-like (Table 2). Previously established gene models in the candidate regions were examined for previously established function using the ensemble Plants BioMart (http://plants.ensembl.org/index.html) and three gene models were selected as candidates (Table 3).

Chromosome	Region	Number of A619 SNP Variants	Number of A632 SNP Variants	Proportion of A619-like SNPs	Proportion of A632-like SNPS
1	1.44e8-1.45e8	10	1	0.91	0.09
1	1.69e8-1.7e8	17	2	0.89	0.11
1	1.87e8-1.88e8	15	0	1	0
2	3.0e6-4.0e6	25	0	1	0
2	1.4e7-1.9e7	309	0	1	0
2	1.9e7-2e7	126	2	0.98	0.02
2	2.0e7-2.1e7	20	0	1	0
2	2.1e7-2.2e7	15	1	0.94	0.06
2	2.2e7-2.3e7	10	0	1	0
2	2.9e7-3e7	29	5	0.85	0.14
2	2.13e8-2.14e8	11	0	1	0
4	3.8e7-3.9e7	12	2	0.86	0.14
4	6.4e7-6.5e7	46	0	1	0
4	6.5e7-6.6e7	29	1	0.97	0.03
4	6.7e7-6.9e7	56	0	1	0
4	7.2e7-7.3e7	17	1	0.94	0.06
4	7.5e7-7.6e7	34	1	0.97	0.03
4	8.2e7-8.3e7	29	2	0.94	0.06
4	9.4e7-9.5e7	10	1	0.91	0.09
6	2.8e7-7.9e7	48	0	1	0
6	1.13e8-1.14	40	5	0.89	0.11
6	1.66e8-1.68e8	28	0	1	0

Table 2: Regions with a high proportion of A619-like SNPs determined by the analysis of the *rmr10-1* RNA-seq data. A high proportion of A619-like character as opposed to A632-like character suggests the region may contain the *rmr10-1* lesion. Coordinates are based on the v4 B73 reference genome.

Gene ID	Coordinates	Gene Description	Predicted SNP Effect
Zm0001d002950	2:27,929,335-2:27,938,858	Protein Chromatin Remodeler	None
Zm00001d003027	**2:29,757,177-2:29,758,837**	**Agenet Domain-Containing /Bromo-Adjacent Homology (BAH) Domain-Containing Protein**	**Missense/Splicing Region**
Zm00001d050149	4:68,410,505-4:68,416,118	Histone Deacetylase	None

Table 3: Three of the candidate genes determined based on their location in regions of interest. These four were selected based on biological functions potentially relevant to epigenetic gene regulation and the *rmr10* mutant phenotype, as well as any SNPs predicted to change protein structure and function. The most likely candidate gene is bolded.

Discussion

The qRT-PCR assay measuring *Pl1-Rh* mRNA abundance provided unexpected results, given the expectation that the *Pl'* state would have a smaller abundance of *Pl1-Rh* mRNA as compared to the *Pl-Rh* state. This assay was limited, however, by the different primer sets used to amplify *Pl1-Rh* mRNA in the A619 control *Pl-Rh* and *Pl'* plants and the *rmr10* mutant and non-mutant siblings. The differences in relative *Pl1-Rh* mRNA abundance could have been due to differential amplification by the two different primer sets. More qRT-PCR data is needed to determine if a loss of RMR10 function leads to increased *Pl1-Rh* mRNA abundance or if RMR10 function represses the production of the anthocyanin pigment independent of the *pl1* locus. All samples should be amplified with the primer used to show a difference between A619 *Pl-Rh* and *Pl'* (pl1_Cone_F and pl1_Cone_R) at the same time to determine if RMR10 has no effect on *Pl1-Rh* mRNA abundance.

If more data confirms these preliminary results, then RMR10, unlike other RMR proteins, is likely not involved in the sRNA biogenesis pathway that is thought to direct epigenetic changes affecting the *Pl1-Rh* allele. The *rmr10* locus may be responsible for gene regulation through a mechanism such as the export, processing, or translatability of *Pl1-Rh* mRNA, or through a mechanism unrelated to *Pl1-Rh* mRNA instead. Given much of our understanding about paramutation is based on this sRNA biogenesis pathway, understanding the function of RMR10 will widen our understanding of how paramutation can occur.

The mapping of the *rmr10* locus using this bioinformatics pipeline provided a small number of candidate regions in the genome and three candidate genes in a fraction of the time compared to traditional mapping techniques. The candidate regions were selected with a reasonable amount of certainty, given the amount of reference sequence (from sources like the HapMap) which are now available. This narrows down the search for the *rmr10* locus from the entire maize genome down to a few select intervals on four chromosomes.

Unfortunately, none of the selected candidate intervals contain gene models that are immediately compelling. The selected candidate gene models (Table 3) are logical choices because their functions could presumably affect gene regulation, but they are either not directly related to previously identified genes implicated in paramutation, do not contain a SNP which is predicted to cause protein disfunction, or do not have adequate read coverage to determine if such a SNP exists.

The first of these genes is a protein chromatin remodeler, a type of protein responsible for affecting the accessibility for DNA to be transcribed.[20] Such an epigenetic function could cause the loss of repression observed in plants homozygous for a mutant *rmr10* allele, but there was no predicted SNP effect in the gene explaining protein disfunction, making it a less likely candidate

than gene models with potentially causative SNP. The gene model does not have complete read coverage in the RNA-seq data set, evident by my visual inspection of the aligned reads using the IGV genome browser (https://software.broadinstitute.org/software/igv/), so the causative lesion may not be represented by the available data. Therefore, genes without potentially causative SNPs cannot be ruled out as candidates yet.

Second was an agenet domain-containing/bromo-adjacent homology (BAH) domain-containing protein. These proteins are thought to help mediate CpG methylation and subsequent gene silencing. Specifically, it is thought that the Dmnt1 protein in mammals, which is responsible for targeting sites of DNA replication and maintaining methylation patterns, contains a BAH domain.[21] This gene model in the *rmr10-1* RNA-seq data set does contain two potentially causative SNP variations, a missense mutation and a mutation in the splicing region. Missense mutations could cause protein disfunction if the mutation causes a non-conservative amino acid substitution. Splicing mutations could be causative if an intron is retained and causes disruption to the amino acid sequence, like the introduction of a stop codon. Given the potentially relevant biological function of this type of protein and its potentially causative SNP variations, this gene model is currently the best candidate for the gene containing the *rmr10-1* lesion.

The third gene encodes a histone deacetylase, a gene responsible for removing acetyl groups from histones and generally promoting inaccessibility and transcriptional deactivation.[22] Such a function could explain the RMR10 mutant phenotype as a loss of histone deacetylase function would cause a loss of transcriptional repression. This locus again has no predicted causative SNP, but sequence coverage is incomplete.

Conclusion

Current data suggests that RMR10 affects anthocyanin production in anthers without affecting *Pl1-Rh* mRNA abundance. More data is needed to evaluate these initial findings; the assay should be repeated with the same primer set used to amplify the A619 controls and under the same experimental conditions. Other assays capable of evaluating how the mutation affects repression of the *Pl1-Rh* allele should also be done, including the analysis of bulk low molecular weight RNA profiles. Since many other *rmr* mutations affect 24nt sRNA biogenesis and sRNAs are thought to be involved in paramutation of the *Pl1-Rh* allele, RMR10 may also influence the accumulation of sRNAs.

Mapping efforts have produced candidate regions and genes that can be further investigated for the *rmr10-1* lesion. Based on the preliminary results, the most plausible candidate gene is an agenet domain-containing/BAH protein which is involved in CpG methylation and gene silencing. This gene is both biologically relevant to epigenetics and the phenomenon of paramutation and has two SNPs that potentially disrupt protein function and cause the loss of gene silencing.

Gene models not present in the RNA-seq data set can also be examined as candidates, as sequences present in the B73 line may be absent in *rmr10-1*. A regulatory mutation in *rmr10-1* may prevent the transcription of a gene model, leading it to be absent in *rmr10-1* but present in the B73 line.

Candidate gene models can be verified in many ways; first, the provisional second allele of *rmr10* (*rmr10-2*) can be verified as an allele of *rmr10* by using the same molecular characterization technique I used to determine *rmr10* and *mop1* were not allelic. Following its molecular characterization, *rmr10-2* can be sequenced to find lesions in the same gene model. If

rmr10-2 contains a lesion in the same gene model as *rmr10-1*, this will further support the identity of *rmr10-2* as an allele of *rmr10*. Continuing allele screens will continue to attempt to generate more mutagenized alleles that can also be sequenced.

Following all forward genetic approaches, the candidate gene can also be disrupted with a variety of reverse genetic techniques. First, there are several maize populations having unique transposon insertions that often disrupt that function of the resident genes. Given a specific candidate gene sequence, like the current candidate gene Zm00001d003027**,** one can request a specific line having such transposon insertions. If such a line is not available for the desired candidate gene or this approach does not disrupt gene function, then gene editing tools can be used to either restore gene function to an otherwise mutant genotype or target the disruption of that gene function and evaluate if such changes phenocopy. This validation will provide conclusive demonstration that I have identified the correct *rmr10* locus.

The identification and characterization of the *rmr10* locus will provide valuable information about how pathways affecting paramutation function mechanistically and provide more information about how gene regulation works both in maize and diverse organisms. Since paramutation involves a complex regulatory pathway, individual components like RMR10 are crucial in understanding the larger pathways and mechanisms. Understanding paramutation will expand our understanding of mechanisms meiotically-heritable changes in gene regulation.

References

[1] Weinhold, B. (2006, March). Epigenetics: The science of change. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1392256/

[2] Brink, R. A. A genetic change associated with the *R* locus in maize which is directed and potentially reversible. *Genetics* **41**, 872–889 (1956).

[3] Hagemann, R. Somatische Konversion bei *Lycopersicon esculentum* Mill. *Z. Vererbungsl.* **89**, 587–613 (1958).

[4] Brink, R. A. Paramutation. *Annu. Rev. Genet.* **7**, 129–152 (1973).

[5] Hollick, J. B. (2016). Paramutation and related phenomena in diverse species. *Nature Reviews Genetics, 18*(1), 5-23. doi:10.1038/nrg.2016.115

[6] Hollick, J.B., & Chandler, V.L. (2001). Genetic factors required to maintain repression of a paramutagenic maize *pl1* allele. *Genetics*, 157 1, 369-78.

[7] Hollick JB, Patterson GI, Coe EH Jr, Cone KC, Chandler VL. Allelic interactions heritably alter the activity of a metastable maize *pl* allele. *Genetics*. 1995;141(2):709–719.

[8] Erhard, K. F., Talbot, J. R., Deans, N. C., Mcclish, A. E., & Hollick, J. B. (2015). Nascent Transcription Affected by RNA Polymerase IV in *Zea mays*. Genetics, 199(4), 1107-1125. doi:10.1534/genetics.115.174714

[9] Nobuta, K., Lu, C., Shrivastava, R., Pillay, M., De Paoli, E., & Accerbi, M. et al. (2008). Distinct size distribution of endogenous siRNAs in maize: Evidence from deep sequencing in the *mop1-1* mutant. *PNAS*, 105(39), 14958-14963. doi:10.1073/pnas.0808066105

[10] Stonaker, J. L., Lim, J. P., Erhard, K. F., Hollick, J. B., & Franz, M. (2009). Diversity of Pol IV Function Is Defined by Mutations at the Maize *rmr7* Locus. *PLoS Genetics*, 5(11), e1000706. https://doi.org/10.1371/journal.pgen.1000706

[11] Erhard, K., Stonaker, J., Parkinson, S., Lim, J., Hale, C., & Hollick, J. (2009). RNA Polymerase IV Functions in Paramutation in *Zea mays*. *Science*, 323(5918), 1201-1205. doi:10.1126/science.1164508

[12] Hale, C. J., Stonaker, J. L., Gross, S. M., & Hollick, J. B. (2007). A Novel Snf2 Protein Maintains trans-Generational Regulatory States Established by Paramutation in Maize. *PLoS Biology*, 5(10). doi:10.1371/journal.pbio.0050275

[13] A. Narain, I. Liao, J. Hollick, unpublished

[14] Barbour, J., Liao, I., Stonaker, J., Lim, J., Lee, C., & Parkinson, S. et al. (2012). required to maintain repression2 Is a Novel Protein That Facilitates Locus-Specific Paramutation in Maize. *The Plant Cell*, 24(5), 1761-1775. doi:10.1105/tpc.112.097618

[15] Hale, C. J., Stonaker, J. L., Gross, S. M., & Hollick, J. B. (2007). A Novel Snf2 Protein Maintains trans-Generational Regulatory States Established by Paramutation in Maize. *PLoS Biology*, 5(10). doi:10.1371/journal.pbio.0050275

[16] Gross, S. M., & Hollick, J. B. (2007). Multiple Trans-Sensing Interactions Affect Meiotically Heritable Epigenetic States at the Maize *pl1* Locus. *Genetics*, 176(2), 829-839. doi:10.1534/genetics.107.072496

[17] J. Hollick, unpublished

[18] Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. "A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3." *Fly* (Austin). 2012 Apr-Jun;6(2):80-92. PMID: 22728672 doi:10.1534/genetics.105.045260

[19] Mike Sovic, Unpublished

[20] Marfella, C. G., & Imbalzano, A. N. (2007). The Chd family of chromatin remodelers. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 618(1-2), 30-40. doi:10.1016/j.mrfmmm.2006.07.012

[21] Callebaut, I., Courvalin, J., & Mornon, J. (1999). The BAH (bromo-adjacent homology) domain: A link between DNA methylation, replication and transcriptional regulation. *FEBS Letters*, 446(1), 189-193. doi:10.1016/s0014-5793(99)00132-5

[22] Xu, W. S., Parmigiani, R. B., & Marks, P. A. (2007). Histone Deacetylase Inhibitors. *Oncogene*. doi:10.1007/springerreference_169351

[23] A Narain, Unpublished

[24] Talbot, J. R. *Maize RNA polymerase IV complexes and their control of gene function* (Doctoral dissertation). University of California, Berkeley.

Supplemental Figures

Primer	Sequence
umc 1845F	**5'-TGGTTGAACTGTTAAATCTGTCCTGA-3'**
umc 1845R	**5'-TGGTAACCAGATTCCCACAGATG-3'**
ANSSR-9F	**5'-TCAGGATTGTGTCTCCATGC-3'**
ANSSR-9R	**5'-GCCGCTAGCAAGCTGTATCT-3'**
ANSSR-10F	**5'-GTCATCAACGCGCTCAACT-3'**
ANSSR-10R	**5'-CCTGCCACGTTTTTCTCTCT-3'**
ASN34F	**5'- CTGCACCGACAAATGAAGAA -3'**
ASN34R	**5'-CTGCTCCTCCTGGACTTGAG-3'**
ANSSR-13F	**5'-TCAATGTTGGTTGATGTTTGG-3'**
ANSSR-13R	**5'-GGAAGAACATGAGTCTTCTTTGC-3'**
ASN41F	**5'-GTCTCCGTAATCGGTCGTTG-3'**
ASN41R	**5'-CGGTCTGCCTTTAGGTGTGT-3'**

Table S1: dCAPS primer sequences used to amplify the area surrounding the *mop1* locus. [23]

PCR Reaction Component	**Amount**
1x PCR mix (lab-made)	14.95 µL
Forward Primer	2 µL
Reverse Primer	2 µL
NEB Thermo Taq Enzyme	0.08 µL
DNA template	1 µL
Thermocycler Setting	**Steps**
55˚ Touchdown	1. 95° for 5 mins 2. 95° for 30s 3. 63° for 30s 4. 72° for 1 min 5. Repeat 2-4 for a total of 16 times, lowering the temperature of step 2 by 1˚ every 2 cycles, until 55˚ is reached. 6. Repeat 2-4 35 times once 55˚ is reached. 7. 72° for 10 mins 8. 10° rest
Restriction Digest Mixture	**Amount**
NEB Cutsmart Buffer	9 µL
DpnI	1.2 µL
H_2O	67.8 µL
PCR Product	4 µL
Incubation Temperature	**Incubation Time**
37˚C	1 hour

Table S2: Specifications of the PCR and restriction digest used to amplify and digest sequence surrounding the *mop1* locus.

Primer	Sequence
pl1_FP2	**5'-GGATCTCATCGTCCGGCTCCACAA -3'**
phi031R	**5'-CCAGCGTGCTGTTCCAGTAGTT-3'**
pl1_Cone_F	**5'-CACGGCGAAGGCAAATGGAG-3'**
pl1_Cone_R	**5'-CTGTTGCCGAGGAGCTTGTG -3'**
alt203_FP1	**5'-CAATATCACTGGTCAAATCCTTGCGA-3'**
aatR	**5'-TTGCACGACGAGCTAAAGACT-3'**

Table S3: Primer sequences to quantify *Pl1-Rh* mRNA abundance in *rmr10* and A619 controls. Pl1_FP2 and phi031R were used to amplify *Pl1-Rh* in *rmr10* samples, pl1_Cone_F and pl1_Cone_R were used to amplify *Pl1-Rh* in A619 control samples, and alt203_FP1 and aatR were used to amplify *aat* in both assays. [23, 24]